



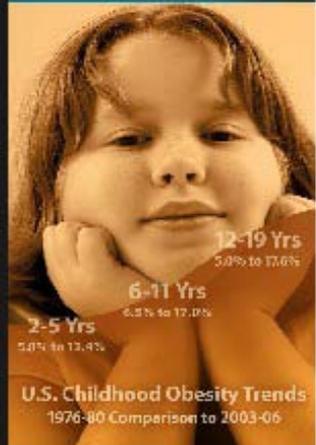
NTP
National Toxicology Program

NTP Workshop: Role of Environmental Chemicals in the Development of Diabetes and Obesity

Breakout Group on Bioinformatics and High Throughput Screening (HTS)

William Janzen (chair)
Vickie Walker (rapporteur)

**Crabtree Marriott Hotel
January 11-13, 2011**



Bioinformatics and HTS Breakout Group Members

Scott Auerbach, NIH/NIEHS/NTP

Chirag Patel, Stanford

Bruce Blumberg, UC-Irvine

David Reif, US EPA

Deborah Clegg, U Texas Southwestern

Jennifer Schlezinger, Boston U

David Dix, US EPA

Supriya Srinivasan , Scripps Research Institute

Alison Holloway, McMaster U

Vickie Walker, NIH/NIEHS/NTP (**rapporteur**)

Keith Houck, US EPA

Morris White, Harvard

William Janzen, UNC – Chapel Hill (**chair**)

Richard Judson, US EPA

James Kaput, FDA/NCTR

Franck Mauvais-Jarvis, Northwestern

Angel Nadal, Miguel Hernandez U

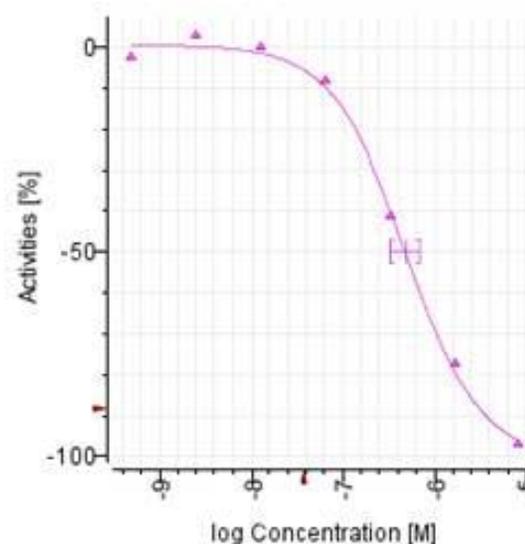


In response to the 2007 NAS report on “Toxicity Testing in the 21st Century: A Vision and a Strategy”, a Memorandum of Understanding was signed in 2008 and expanded in 2010 to include the FDA

- **National Toxicology Program:** Linda S. Birnbaum, Ph.D., DABT, ATS Director, National Institute of Environmental Health Sciences, National Institutes of Health
- **NIH Chemical Genomics Center:** Eric D. Green, M.D., Ph.D. Director National Human Genome Research Institute
- **U.S. Environmental Protection Agency:** Paul T. Anastas, Ph.D. Assistant Administrator Office of Research and Development
- **Food and Drug Administration:** Janet Woodcock, M.D., Director Center for Drug Evaluation and Research

The NCGC

- Conducts quantitative high-throughput screening (qHTS)
 - >300,000 profiles/week
- qHTS profile
 - 1536-well plate format
 - 15-point concentration-response curve
 - DMSO soluble
 - 5 nM to 92 μ M typical
 - ~5 μ L assay volume
 - ~1000 cells/well



NCGC: Phase I qHTS screening

- ~2800 compounds provided by the NTP and the EPA

Apoptosis

- 3/7, 8, & 9

Cell Viability

- ATP
- LDH
- Protease release

DNA damage

- ELG1
- p53
- Multiple repair-

deficient DT40
cell lines

Epigenetics

- LDR

Mitochondrial toxicity

Nuclear Receptors

- hAR
- hER α
- hFXR
- hGR

- hLXR β

- hPPAR α

- hPPAR γ

- hPPAR δ

- hPXR

- rPXR

- hROR

- hRXR

- hTR β

- hVDR

Pathways

- AP1

- ARE/Nrf2

- CRE

- hERG

- HRE

- HSP

- JNK3

- NF κ B

NCGC: Phase II

- Will be screening >10,000 compounds, with 3X concentration response curves
- Initial screening focus on nuclear receptors and stress response pathways; then will focus on disease oriented pathways

NCGC

- Drugs
- Drug-like compounds
- Active pharmaceutical ingredients

EPA

- ToxCast I and II compounds
- Antimicrobial Registration Program
- Endocrine Disruptor Screening Program
- OECD Molecular Screening Working Group List
- Failed Drugs

NTP

- NTP-studied compounds of all types
- NTP nominations and related compounds
- ICCVAM and NICEATM validation and reference compounds
- Outside collaborators (e.g., U.S. Army Public Health Command)



EPA's ToxCast™ Program

- Research program of EPA's National Center for Computational Toxicology (NCCT) (see <http://www.epa.gov/ncct/toxcast>)
- Addresses chemical screening and prioritization needs for EPA
- Comprehensive use of HTS technologies to generate biological fingerprints and predictive signatures
- Data released via ACToR (Aggregated Computational Toxicology Resource) (<http://epa.gov/actor>)
 - Contains data on ~500,000 environmental chemicals
 - Multiple Domains - Physchem, biological, use levels, regulations
 - Brings together data from >200 sources

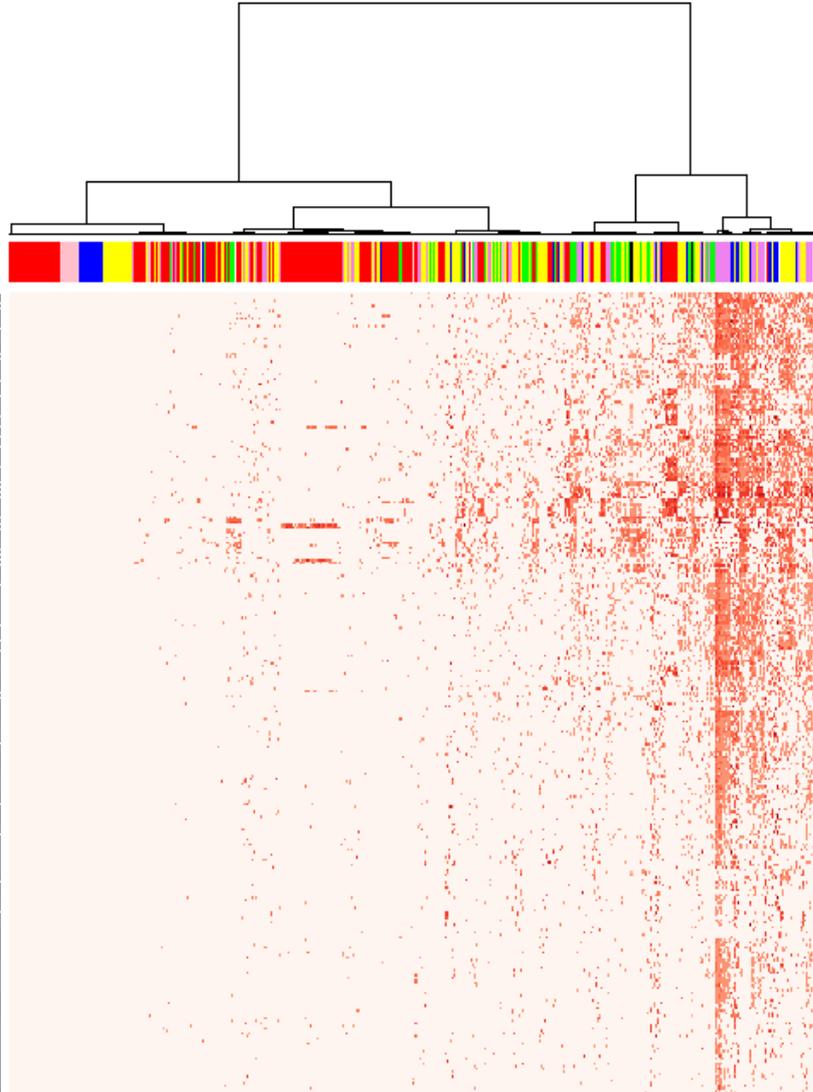
Phased Development of ToxCast™

Phase	Number of Chemicals	Chemical Criteria	Purpose	Number of Assays	Target Date
<i>I</i>	320	<i>Data Rich (pesticides)</i>	<i>Signature Development</i>	554	<i>FY08</i>
Ib	15	Nanomaterials	Pilot	166	FY09
IIa	>300	Data Rich Chemicals	Validation	>400	FY10
IIb	>100	Known Human Toxicants	Extrapolation	>400	FY10
IIc	>300	Expanded Structure and Use Diversity	Extension	>400	FY10
IIId	>12	Nanomaterials	PMN	>200	FY10
IIe	50 NTP	Immunotoxicants	Signature development	>400	FY10
III	Thousands	Data poor	Prediction and Prioritization	>300	FY 12

The ToxCast™ In Vitro Data Set

Assays →

Chemicals ↓



Many hits – median value=50

Fewer in cell-free HTS (Novascreen, red)

Many hits are at or near top of tested concentration range

A few are in nanomolar range



The NCGC BioPlanet

- Hosts the universe of pathways
- All pathway annotations from manually curated, public sources
- Integrates pathways from different data sources
- Annotates pathways by source, species, biological function/process, disease/toxicity relevance, assay availability
- Easy visualization, browsing, analysis of pathways
- Facilitates pathway assay selection/prioritization for Tox21 production phase
- Will be publicly available

The Universe of Pathways

Detailed view of a pathway

The screenshot displays the NOGIC DiaPlanet interface. On the left, a search panel includes fields for 'Category', 'GeneKeyword', and 'Find', along with filters for 'Toxicity', 'Disease', and 'PubChem'. Below the search panel is a large blue 'Pathways' globe. On the right, a detailed view of a pathway is shown as a network of nodes and edges. Below this, a 'Find Gene' window displays information for the 'erk1/erk2 mapk signaling pathway' (ID: 100170).

Gene ID	Gene Symbol	Gene Description
9252	RPS6KA5 RLPK MGC1911 MSPK1 MSK1	ribosomal protein S6 kinase, 90kDa, polypeptide 5
6714	SRC1 ASV p60-Src c-SRC SRC	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
8569	MNK1 MNKN1	MAP kinase interacting serine/threonine kinase 1
6195	IRU-1 RPS6KA1 RSK1 RSK MAPKAPK1A	ribosomal protein S6 kinase, 90kDa, polypeptide 1
5595	HS44KDAP ERK1 P44MAPK HUMKER1A P44ERK1 MAPK3 MGC20180 PRKM3	mitogen-activated protein kinase 3
2872	MEK2 GPRK7 MNK2	MAP kinase interacting serine/threonine kinase 2
5605	MEK2 MAP2K2 MAPK32 PRKME2 MEK2 FLJ26075	mitogen-activated protein kinase kinase 2
5528	MGC8949 PPP2R5D B56D MGC2134	protein phosphatase 2, regulatory subunit B', delta isoform
5894	RAF1 c-Raf Raf-1 NS5 CRAF	v-raf-1 murine leukemia viral oncogene homolog 1
4803	HSAN5 Beta-NGF MGC161426 MGC161428 NGFB NGF	nerve growth factor (beta polypeptide)
2002	ELK1	ELK1, member of ETS oncogene family
5801	FLJ34328 MGC148170 PCPTP1 PTPBR7 PTPRR DKFZp781C1038 PTP-SL MGC131968 PTPRQ EC-PTP	protein tyrosine phosphatase, receptor type, F
4804	CD271 Gp80-LNGFR p75(NTR) TNFRSF16 NGFR p75NTR	nerve growth factor receptor (TNFR superfamily, member 16)
5594	MAPK2 MAPK1 p41mapk ERK2 p38 ERK P42MAPK p40 p41 PRKM2 PRKMI ERT1	mitogen-activated protein kinase 1
6774	FLJ20882 IHES MGC16063 APRF STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)
7015	TRT EST2 hEST2 TCS1 TP2 TERT	telomerase reverse transcriptase
3265	H-RASIDK N-RAS HRAS K-RAS C-HA-RAS1 C-BAS/MS RASH1 C-H-RAS HAMS/ CTLO HRAS1	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
5604	MEK1 MAP2K1 MAPK31 PRKMK1 MEK1	mitogen-activated protein kinase kinase 1
4609	hHLHe39 MRTL MYC c-Myc	v-myc myelocytomatosis viral oncogene homolog (avian)

Pathways

Gene information

~1100 human pathways mapped to the pathway globe

Pathways containing obesity genes

The NCGC BioPlanet
[Close] [Maximize]

Category: Toxicity Disease PubChem

Gene: 208 359 951 554

Find

And Or

0 25 50 75 100

229 pathways found.

Find Gene:

Information

hsa04930: Type II diabetes mellitus

Gene ID	Gene Symbol	Gene Description
5602	PRKM10 MGC50974 JNK3A FLJ33785 MAPK10 FLJ12099 p54bSAPK JNK3 p493F12	mitogen-activated protein kinase 10
5291	PIK3CB PI3KBETA DKFZp779K1237 PIK3C1 PI3K P110BETA MGC133043	phosphoinositide-3-kinase, catalytic, beta polypeptide
6517	SLC2A4 GLUT4	solute carrier family 2 (facilitated glucose transporter), member 4
6833	TNDM2 ABC36 SUR HI MRP8 HRINS SUR1 HHF1 PHH1 ABCC8	ATP-binding cassette, sub-family C (CFTR/MRP), member 8
5296	P85B p85 p85-BETA PIK3R2	phosphoinositide-3-kinase, regulatory subunit 2 (beta)
5581	nPKC-epsilon PRKCE MGC125656 PKCE MGC125657	protein kinase C, epsilon
5293	P110DELTA PIK3CD p110D PI3K	phosphoinositide-3-kinase, catalytic, delta polypeptide
5594	MAPK2 MAPK1 p41mapk ERK2 p38 ERK P42MAPK p40 p41 PRKM2 PRKM1 ERT1	mitogen-activated protein kinase 1
776	Cav1.3 CACN4 CACH3 CACNL1A2 CCHL1A2 CACNA1D	calcium channel, voltage-dependent, L type, alpha 1D subunit
8471	IRS4 IRS-4 PY160	insulin receptor substrate 4
3098	HK1 HK1 HXK1 HK1-tb HK1-ta HK1-tc	hexokinase 1
2475	FLJ44809 FRAP2 FRAP1 RAPT1 FRAP MTOR RAFT1	mechanistic target of rapamycin (serine/threonine kinase)
5599	JNK PRKM8 JNK1A2 JNK1 JNK21B1/2 MAPK8 SAPK1	mitogen-activated protein kinase 8
7124	TNFA TNF TNFSF2 DIF TNF-alpha	tumor necrosis factor (TNF superfamily, member 2)
5295	p85-ALPHA GRB1 p85 PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)
3101	HXK3 HK3 HKIII	hexokinase 3 (white cell)
777	CACNL1A6 CACNAIE CACH6 Cav2.3 BII	calcium channel, voltage-dependent, R type, alpha 1E subunit
3630	IRDN INS ILPR	insulin
8835	SOCS-2 Cish2 CIS2 SOCS2 SSI2 STAT12 SSI-2	suppressor of cytokine signaling 2
23533	FOAP-2 p101 PIK3R5 F730038115Rik P101-PI3K	phosphoinositide-3-kinase, regulatory subunit 5
389692	hMAFA RIPE3b1 MAFA	v-maf musculoaponeurotic fibrosarcoma oncogene homolog A (avian)

12

Pathways containing obesity genes

The NCGC BioPlanet
[Close] [Maximize]

Category: Toxicity Disease PubChem

Gene: 154 155 181 1009

Find:

And Or

71 pathways found.

Find Gene:

Information

100004: visceral fat deposits and the metabolic syndrome

Gene ID	Gene Symbol	Gene Description
2908	GRL GR GCCR NR3C1 GCR	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)
7124	TNFA TNF TNFSF2 DIF TNF-alpha	tumor necrosis factor (TNF superfamily, member 2)
56729	RETN ADFR RSTN MGC126609 RETN1 FIZZ3 MGC126603 XCP1	resistin
5468	GLM1 NR1C3 PPARC PPARG2 PPARgamma PPARG1 CIMT1	peroxisome proliferator-activated receptor gamma
6256	FLJ16733 FLJ00280 MGC102720 FLJ16020 RXRA NR2B1 FLJ00318	retinoid X receptor, alpha
4023	LPL HDLCQ11 LIPD	lipoprotein lipase
9370	GBP28 adiponectin ACDC ADIPQTL1 ADPN APM-1 ACRP30 ADIPOQ APM1	adiponectin, C1Q and collagen domain containing
3290	MGC13539 HDL 11-DH HSD11B 11-beta-HSD1 HSD11 HSD11B1 HSD11L SDR26C1	hydroxysteroid (11-beta) dehydrogenase 1



NTP
National Toxicology Program

Data Resources



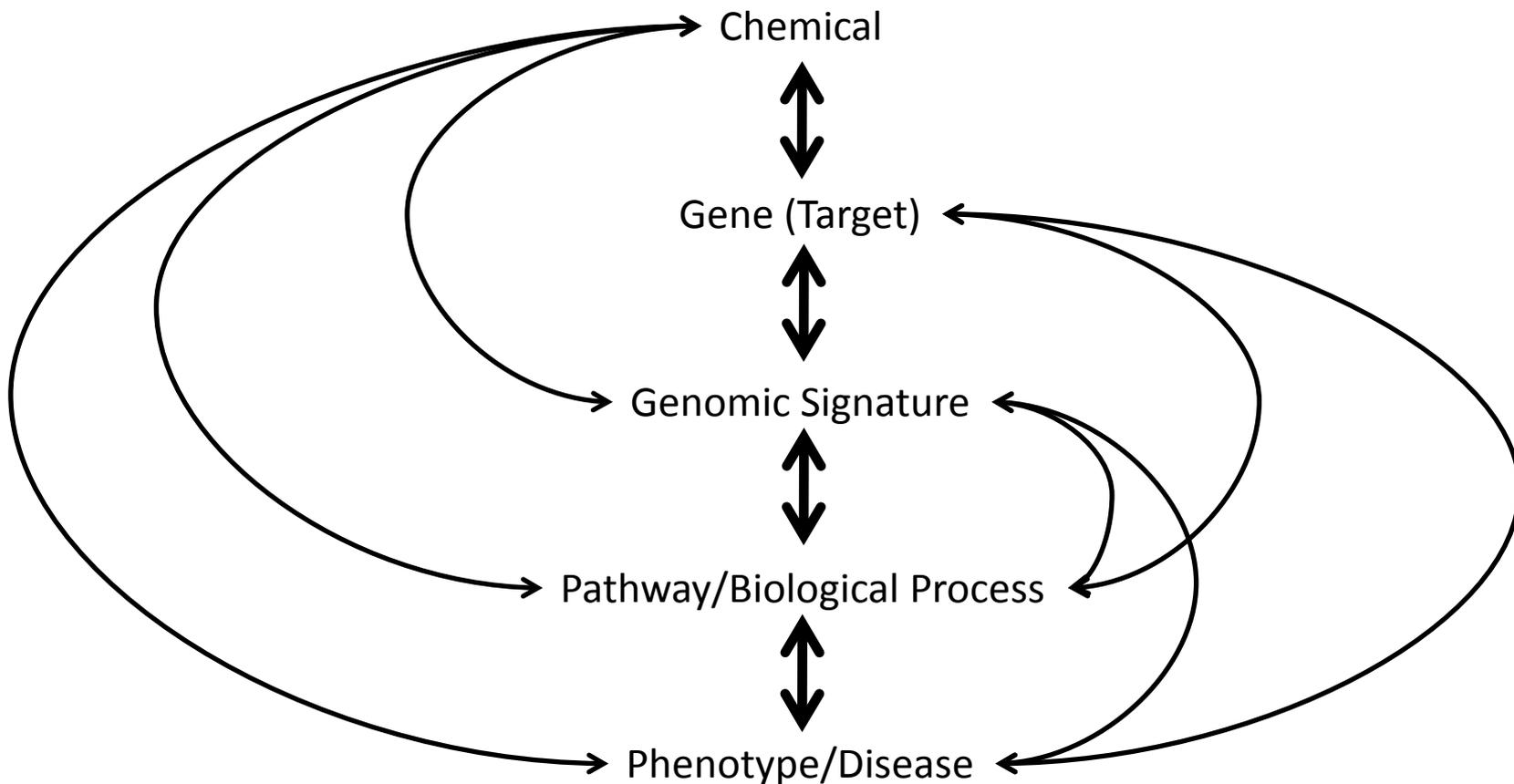


Tox21 Data

- ToxCastDB contains the ToxCast and ToxRefDB data for browsing
 - <http://actor.epa.gov/actor/faces/ToxMiner/Home.jsp>
- ACToR contains a much large amount of data on other types of toxicity and exposure
 - <http://actor.epa.gov>
- NTP CEBS (Chemical Effects in Biological Systems) – NTP historical data, DrugMatrix, HTS, *C. elegans*, etc.
 - <http://www.niehs.nih.gov/research/resources/databases/cebs/>
- PubChem
 - <http://pubchem.ncbi.nlm.nih.gov/>



Relationships Abound





Database Resources: Chemicals to Genes (Targets)

- DrugBank
 - <http://www.drugbank.ca/>
- FooDB
 - <http://www.drugbank.ca/>
- Small Molecule Pathway Database (SMPDB)
 - <http://www.smpdb.ca/>
- Toxin and Toxin Target Database
 - <http://www.t3db.org/>
- Potential Drug Target Database
 - <http://www.dddc.ac.cn/pdtd/>
- KEGG Compound (Ligand)
 - <http://www.genome.jp/kegg/compound/>
 - <http://www.genome.jp/ligand/>
- STITCH
 - <http://stitch.embl.de/>
- PharmGKB
 - <http://www.pharmgkb.org/>
- PubChem
 - <http://pubchem.ncbi.nlm.nih.gov/>
- SuperTarget
 - <http://bioinf-tomcat.charite.de/supertarget/>
- SuperToxic
 - <http://bioinf-services.charite.de/supertoxic/>
- ChEBML
 - <https://www.ebi.ac.uk/chembl/db/index.php>
- Noted earlier: Ingenuity, GeneGo CoPub, CTD
- PubChem
 - <http://pubchem.ncbi.nlm.nih.gov/>
- SuperTarget
 - <http://bioinf-tomcat.charite.de/supertarget/>
- ChEBML
 - <https://www.ebi.ac.uk/chembl/db/index.php>
- GeneGo
 - Hand curated database with variety of gene ontologies
 - <http://www.genego.com/>
- Ingenuity
 - Hand curated database with variety of gene ontologies
 - <http://www.ingenuity.com/>
- NextBio
 - Integrated public data mining resource for transcriptomics, CGH, genome-wide methylation
 - <http://www.nextbio.com/b/nextbio.nb>
- Omic Browser
 - <http://omicspace.riken.jp/>
- Comparative Toxicogenomics Database (CTD)
 - Text mining (hand curation)
 - <http://ctd.mdibl.org/>
- Ki Database
 - <http://pdsp.med.unc.edu/kidb.php>

Database Resource: Chemicals to Genomic Signatures

- ArrayExpress
 - <http://www.ebi.ac.uk/arrayexpress/>
- Connectivity Map
 - http://www.broadinstitute.org/genome_bio/connectivitymap.html
- DrugMatrix (freely available through NIEHS by mid 2011)
- GEO
 - <http://www.ncbi.nlm.nih.gov/geo/>
- NextBio
 - <http://www.nextbio.com/b/nextbio.nb>



Database Resource: Chemicals to Pathways

- CoPub
 - <http://services.nbic.nl/copub/portal/>
- Comparative Toxicogenomics Database (CTD)
 - <http://ctd.mdibl.org/>
- GeneGo
 - <http://www.genego.com/>
- Ingenuity
 - <http://www.ingenuity.com/>
- KEGG Compound/Ligand
 - <http://www.genome.jp/kegg/compound/>
 - <http://www.genome.jp/ligand/>
- Omic Browser
 - <http://omicspace.riken.jp/>

Database Resources: Chemicals to Phenotype/Disease

- Side Effect Resource (SIDER)
 - <http://sideeffects.embl.de/>
- Drug Adverse Reaction Database (DART)
 - <http://xin.cz3.nus.edu.sg/group/drt/dart.asp>
- NTP Database
 - <http://ntp.niehs.nih.gov/?objectid=72016020-BDB7-CEBA-F3E5A7965617C1C1>
- ToxRefDB
 - <http://actor.epa.gov/toxrefdb/faces/Home.jsp>
- Carcinogen Potency Database
 - <http://potency.berkeley.edu/>
- REPDOSE
 - <http://www.fraunhofer-repdose.de/>
- CHE Toxicant and Disease Database
 - <http://www.healthandenvironment.org/tddb>
- IARC Monographs
 - <http://monographs.iarc.fr/>
- ToxNet
 - <http://toxnet.nlm.nih.gov/index.html>
- ACToR
 - <http://actor.epa.gov/actor/faces/ACToRHome.jsp>
- DSSTox
 - <http://www.epa.gov/ncct/dsstox/>
- ATSDR - Health Effects of Exposure to Substances and Carcinogens
 - <http://www.atsdr.cdc.gov/substances/ToxOrganSystem.asp>
- Comparative Toxicogenomics DataBase
 - <http://ctd.mdibl.org/>
- Omic Browser
 - <http://omicspace.riken.jp/>
- eChemPortal
 - http://www.echemportal.org/echemportal/index?pageID=0&request_locale=en

Data Resources: Genes (Targets) to Genomic Signatures

- NextBio
- DrugMatrix (freely available through NIEHS by mid 2011)
- Connectivity Map
 - http://www.broadinstitute.org/genome_bio/connectivitymap.html
- ArrayExpress
 - <http://www.ebi.ac.uk/arrayexpress/>
- GEO
 - <http://www.ncbi.nlm.nih.gov/geo/>



Database Resources: Genes (Targets) to Pathways & Biological Processes

- CoPub
 - <http://services.nbic.nl/copub/portal/>
- Comparative Toxicogenomics Database (CTD)
 - <http://ctd.mdibl.org/>
- GeneGo
 - <http://www.genego.com/>
- Ingenuity
 - <http://www.ingenuity.com/>
- NextBio
 - <http://www.nextbio.com/b/nextbio.nb>
- Omic Browser
 - <http://omicspace.riken.jp/>

Database Resources: Genes (Targets) to Disease

- Catalog of Published Genome-Wide Association Studies
 - <http://www.genome.gov/gwastudies/#1>
- CoPub
 - <http://services.nbic.nl/copub/portal/>
- Comparative Toxicogenomics Database (CTD)
 - <http://ctd.mdibl.org/>
- dbGAP
 - <http://www.ncbi.nlm.nih.gov/gap>
- Entrez Gene
 - <http://www.ncbi.nlm.nih.gov/gene>
- GeneCards
 - <http://www.genecards.org/>
- GeneGo
 - <http://www.genego.com/>
- Human Gene Mutation Database
 - <http://www.hgmd.cf.ac.uk/ac/index.php>
- Ingenuity
 - <http://www.ingenuity.com/>
- NextBio
 - <http://www.nextbio.com/b/nextbio.nb>
- OMIM
 - <http://www.ncbi.nlm.nih.gov/omim>
- Omic Browser
 - <http://omicspace.riken.jp/>
- Phenopedia
 - <http://hugenavigator.net/HuGENavigator/startPagePhenoPedia.do>



Database Resources: Genomic Signatures to Pathways

- DAVID Bioinformatics
 - <http://david.abcc.ncifcrf.gov/>
- ToppGene
 - <http://toppgene.cchmc.org/>
- CoPub
 - <http://services.nbic.nl/copub/portal/>
- GeneGo
 - <http://www.genego.com/>
- Ingenuity
 - <http://www.ingenuity.com/>
- GSEA
 - <http://www.broadinstitute.org/gsea/index.jsp>

Database Resources: Genomic Signatures to Disease

- CoPub
 - <http://services.nbic.nl/copub/portal/>
- Comparative Toxicogenomics Database (CTD)
 - <http://ctd.mdibl.org/>
- GeneGo
 - <http://www.genego.com/>
- Ingenuity
 - <http://www.ingenuity.com/>
- NextBio
 - <http://www.nextbio.com/b/nextbio.nb>
- Omic Browser
 - <http://omicspace.riken.jp/>



Database Resources: Pathways to Disease

- CoPub
 - <http://services.nbic.nl/copub/portal/>
- Comparative Toxicogenomics Database (CTD)
 - <http://ctd.mdibl.org/>
- GeneGo
 - <http://www.genego.com/>
- Ingenuity
 - <http://www.ingenuity.com/>
- NextBio
 - <http://www.nextbio.com/b/nextbio.nb>
- Omic Browser
 - <http://omicspace.riken.jp/>



Database Resources: Chemicals to Genomic Signatures

- NextBio
- DrugMatrix (freely available through NIEHS by mid 2011)
- Connectivity Map
 - http://www.broadinstitute.org/genome_bio/connectivitymap.html
- ArrayExpress
 - <http://www.ebi.ac.uk/arrayexpress/>
- GEO
 - <http://www.ncbi.nlm.nih.gov/geo/>



Discussion Points & Recommendations

- Use *in vitro* assays to prototype the process instead of expansively looking at gene disease relationships - go with what we know
- Other methods for approaching identification of assays
 - Identify compounds with known molecular pharmacology (e.g. weight gain)
 - Identify receptor or gene for further screening



Discussion Points & Recommendations

- Proceed With Caution - Can we use a chemical biology approach to predict an effect? More to it than *in vitro* screening
- Bioinformatics approach is not complete
- Cellular phenotype assays rather than gene centric